Reliability-Aware Adaptive Power Allocation in 5G RAN for uRLLC Applications

Shameer Bashir and Samaresh Bera, *Senior Member, IEEE*
Department of Computer Science and Engineering
Indian Institute of Technology Jammu
Jammu and Kashmir, India - 181221
Email: 2024rcs1022@iitjammu.ac.in, s.bera.1989@ieee.org

***Abstract*—Transmission impairments are inherent to cellular communication systems and significantly impact the signal-to-interference-plus-noise ratio (SINR) at the receiver. Although several methods exist to mitigate these impairments, the diverse and stringent quality-of-service (QoS) requirements of 5G and beyond networks make this problem increasingly challenging. Among emerging applications, ultra-reliable and low-latency communications (uRLLC) are particularly demanding due to their strict reliability and latency constraints.**

This paper addresses the adaptive power allocation problem for downlink transmission to satisfy the reliability requirements of uRLLC users. A detailed mathematical model is developed for SINR and block error rate (BLER) between a base station (gNB) and user equipment (UE). The power allocation problem is formulated as a non-convex optimization problem, aiming to allocate optimal transmit power across gNB–UE pairs while maintaining BLER-based reliability targets. To obtain an efficient solution, a Q-learning–based adaptive power allocation algorithm is proposed, offering low computational complexity and near-optimal performance. Simulation results demonstrate that the proposed approach significantly outperforms fixed power allocation schemes, achieving higher reliability and improved BLER retainability in downlink transmission.

***Index Terms*—5G RAN, Optimization, Reinforcement learning, Q-learning, Reliability**

I. Introduction

Fifth-generation (5G) and beyond networks are designed to support applications with diverse Quality-of-Service (QoS) requirements beyond traditional cellular communications [1]. These applications are broadly classified into three categories: enhanced Mobile Broadband (eMBB), which demands high throughput; ultra-Reliable and Low-Latency Communications (uRLLC), which requires stringent reliability and minimal latency; and massive Machine-Type Communications (mMTC), which supports a large number of connected devices.

This work focuses on radio resource allocation in the 5G Radio Access Network (RAN) for uRLLC applications, which require extremely high reliability (up to 99.999%) and very low latency (up to 10 ms) [1], [2]. Meeting both these requirements simultaneously is inherently challenging due to their interdependence. Reliability can often be improved through re-transmissions, but re-transmissions introduce additional latency. Therefore, to meet uRLLC performance targets, the transmitted information must be correctly received on the first attempt.

In practice, wireless transmissions are subject to various impairments, such as faulty feeders, poor voltage standing wave ratio (VSWR), or base-station failures [3]. Several existing studies have proposed methods—ranging from power control to radio resource allocation—to mitigate such impairments and enhance link reliability [4], [5]. However, most of these approaches assume a fixed block error rate (BLER) (typically 10^{-1}) for determining the threshold signal-to-interference-plus-noise ratio (SINR) used in selecting modulation and coding schemes (MCS) [2]. While suitable for eMBB or general traffic, this assumption fails to account for the diverse reliability requirements of emerging uRLLC services.

To address this limitation, this paper revisits the adaptive power allocation problem for 5G and beyond networks, focusing on uRLLC services with heterogeneous reliability targets. Since transmission reliability is inversely related to BLER, adaptive power control becomes essential to maintain service-specific QoS. The key contributions of this work are summarized as follows:

- Mathematical modeling of SINR and BLER for adaptive downlink power allocation considering reliability requirements.
- Formulation of an optimization problem for adaptive and optimal transmit power allocation under constraints such as target BLER and maximum transmission power.
- Development of a reinforcement learning–based (Q-learning) algorithm to solve the non-convex optimization problem efficiently in polynomial time, enabling adaptive power control for each gNB–UE pair in downlink transmission.
- Simulation of the 5G downlink channel for various Modulation and Coding Schemes (MCSs) using MATLAB to derive curve-fit parameters for approximating BLER as a function of SINR. Performance is further evaluated through event-driven simulations, demonstrating that the proposed approach significantly outperforms fixed power allocation in achieving higher reliability and lower BLER.

The rest of the paper is organized as follows. Section II discusses the existing works that focused on adaptive power allocation in cellular networks. Section III presents the detailed network model with optimization problem. Section IV presents the proposed Q-learning-based adaptive power allocation approach. The performance of the proposed approach is presented in Section V. Finally, Section VI concludes the paper while highlighting future research directions.

II. Related Work

In this section, we review existing studies on resource allocation in 5G Radio Access Networks (RANs), with particular emphasis on power allocation at gNBs for downlink transmission [4]–[11].

Mismar et al. [4] proposed a reinforcement learning (RL)–based power allocation method for downlink transmission to achieve target signal-to-interference-plus-noise ratios (SINRs) for cellular users. The authors examined both indoor and outdoor scenarios. In the indoor case, Q-learning was employed to dynamically adjust gNB transmit power to meet target SINRs. For outdoor environments, they considered fault-induced transmission impairments and introduced a self-organizing network fault management framework using deep Q-learning. In a related study, Mismar and Evans [5] investigated the indoor scenario exclusively, proposing a Q-learning-based closed-loop power control scheme to ensure target SINR satisfaction for cellular users.

Nguyen et al. [6] developed an adaptive power allocation scheme for multi-connectivity MIMO systems, targeting quality-of-experience (QoE) enhancement under imperfect channel state information. Their method leveraged multi-connectivity to improve reliability for critical users. In [7], the authors formulated the adaptive power allocation problem as a sum-rate maximization task with fairness constraints to guarantee a minimum data rate for each user. The problem was solved using a graph neural network–based approach.

Rajanandini and Jaya [11] proposed a meta-heuristic-based adaptive power allocation strategy considering energy efficiency. Rabee et al. [8] investigated throughput-optimized power allocation in energy-harvesting relay-assisted networks, employing an actor–critic RL framework to achieve polynomial-time convergence. Tran et al. [9] examined BLER-based power allocation for mMTC applications. However, their model lacked power allocation constraints, which is impractical, and assumed a system-wide BLER target rather than user-specific reliability requirements. Gao et al. [10] applied Q-learning for power control to improve throughput, energy efficiency, and user experience.

While these studies have explored adaptive power allocation in cellular networks, most have not explicitly addressed reliability, which is crucial for mission-critical uRLLC applications. Among the closest works, reliability-aware multi-connectivity [6] and BLER-based power allocation [9] provide valuable insights but also exhibit key limitations. First, multi-connectivity may be impractical in dense deployments due to limited radio resources and infrastructure constraints. Second, meta-heuristic approaches such as those in [9] lack convergence guarantees, making them unsuitable for near real-time decision-making required in dynamic 5G environments.

III. System Model

We consider a network consisting of multiple gNBs and multiple users, denoted by a set $\mathcal{G}$ and N_{UE}, respectively. The gNB sends information to the users using downlink. In order to receive the information correctly, the signal-to-interference-plus-noise-ratio (SINR) needs to be above the threshold for a specific modulation and coding scheme [12]. Furthermore, individual users may have different reliability requirements in terms of BLER based on the underlying applications, as discussed in Section I. We present the SINR and BLER models in the subsequent sections.

A. SINR and BLER Model

The received base-band signal $\mathrm{y}_i[t]$ for user i at time t is denoted by:

$$\mathrm{y}_i[t] = \mathrm{h}_i[t]\mathrm{x}_i[t] + \eta_0[t], \forall i \in N_{\text{UE}}, \tag{1}$$

where, $\mathrm{h}_i[t]$ is the channel coefficient that depends on attenuation and delay factors. $\eta_0[t]$ is the Gaussian noise sampled from $\mathcal{N}(0, \sigma_n^2)$.

For a given forward link budget, the gNB allocates a transmit power for user i at time t. Therefore, the received power $P_{\text{UE}}^{(i)}[t]$ (in dB) is calculated as follows:

$$P_{\text{UE}}^{(i)}[t] = P_{\text{tx}}^{(i)}[t] + G_{\text{tx}} - L_m - L_p^{(i)}[t] + G_{\text{rcv}}^i, \forall i \in N_{\text{UE}},$$

where, $P_{\text{tx}}^{(i)}$ denotes the transmit power allocated to user i by the transmitting gNB. G_{TX} denotes the antenna gain of the transmitter. L_{m} represents the loss due to transmission impairments, discussed later. $L_{\text{P}}^{(i)}$ denotes the path loss and G_{rcv}^i denotes the antenna gain of the receiving UE.

The received downlink SINR $\gamma^{(i)}$ for the i-th UE at TTI t is represented as:

$$\gamma^{(i)} \triangleq \frac{P_{\text{UE}}^{(i)}}{\eta_0 + \sum\limits_{j \in \mathcal{G} \setminus \{0\}} P_{\text{UE},O_j \to i}}, \forall i \in N_{\text{UE}}. \tag{2}$$

For ease of notations, we dropped the time index t. In (2), η_0 denotes the noise for the AWGN channel. $P_{\text{UE},O_j \to i}$ denotes the inter-cell inference to the user i from all gNBs $j \in N_{\text{gNB}}$ except the gNB to which the user is associated.

Now, the BLER for the transmission with SINR $\gamma^{(i)}$ can be accurately approximated using the following expression [12]:

$$\mu[\gamma^{(i)}] \approx \begin{cases} 1, & \text{if } 0 < \gamma^{(i)} \leq \gamma_{th}, \\ C_m \exp(-d_m \gamma^{(i)}), & \text{if } \gamma^{(i)} > \gamma_{th}, \end{cases} \tag{3}$$

where, γ_{th} denotes the threshold SINR, which depends on the modulation and coding used for the transmission [2]. C_m and d_m denote the constant values associated with modulation and coding [12].

Putting (2) and (3) together, we get

$$\mu^{(i)} = C_m \exp\Big(- d_m \frac{P_{\text{UE}}^{(i)}}{\eta_0 + \sum\limits_{j \in \mathcal{G} \setminus \{0\}} P_{\text{UE},O_j \to i}}\Big), \tag{4}$$

which is equivalent to

$$P_{\text{UE}}^{(i)} = -\frac{1}{d_m}\left[\eta_0 + \sum_{j\in\mathcal{G}\setminus\{0\}} P_{\text{UE},O_j\to i}\right]\log\left(\frac{\mu^{(i)}}{C_m}\right) \quad (5)$$

Finally, we present the expression of the transmit power for the UE i as follows:

$$P_{\text{tx}}^{(i)}[t] = L_m^{(i)} + L_p^{(i)}[t] - G_{\text{UE}} - G_{\text{tx}} - \frac{1}{d_m}\left[\eta_0 + \sum_{j\in\mathcal{G}\setminus\{0\}} P_{\text{UE},O_j\to i}\right]\log\left(\frac{\mu^{(i)}}{C_m}\right). \quad (6)$$

B. Problem Statement

The objective is to adjust the transmit power for each user at TTI t, as presented in (6), to meet the target BLER for reliable transmission. Mathematically,

$$\text{Minimize} \sum_{t=1}^{\tau}\sum_{i=1}^{|N_{\text{UE}}|} P_{\text{tx}}^{(i)}[t], \quad (7)$$

subject to

$$\mu^{(i)}[t] \le \mu_{\text{threshold}^{(i)}}, \forall i \in N_{\text{UE}}, \quad (8a)$$

$$P_{\text{tx}}^{(i)}[t] \le P_{\text{tx}}^{\max}, \forall t \in \{1,\ldots,\tau\}, \quad (8b)$$

where, (7) denotes that our objective is to minimize the total transmission power for all users over a period of transmission. Equation (8a) ensures that the achieved BLER is always less than the BLER threshold for meeting the reliability of the associated application. Equation (8b) denotes the total transmission power capacity constraint. The above optimization problem needs to be run for each gNB-UE pair. Therefore, in a large-scale networks with multiple gNBs and UEs, the problem becomes combinatorial. Consequently, finding optimal solution to the problem is infeasible in polynomial time. Furthermore, the problem becomes more complex in the presence of transmission impairments, which are non-deterministic in nature. We consider the following transmission impairments: faulty feeder, improper voltage standing wave ratio (VSWR), and base-station down. Therefore, the network events related to a downlink transmission are as follows: faulty feeder, improper VSWR, base-station down, feeder is restored, VSWR restored, base-station up. We note that the restore operations are always followed by the associated problem, without which restore operations cannot take place. To solve the problem in polynomial time, we apply Q-learning, which is discussed in subsequent sections.

IV. Proposed Approach: Q-Learning-based Power Allocation

A. Overview of Q-Learning and Associated Parameters

The Q-learning depends on the Markov decision process (MDP) [13], which includes a set of states, a set of actions, transition probabilities, and rewards, which are discussed below.

- *Actions*: The Q-learning agent, called RL-agent, tunes the transmit power for each user by issuing power control commands as actions. We outline the set of actions, denoted by $\mathcal{A}$, on power control commands for a user in Table I. The symbol PC$[t]$ denotes the power control command at time t. The symbol κ_t represents the number of times the PC$[t]$ command is executed at time t.

TABLE I: Set of actions for transmit power control

Action	Power control (in dB)
0	PC$[t] = 0$: No change in transmit power
1	PC$[t] = -0.1$: Decrease the transmit power thrice, i.e., total change is -0.3, ($\kappa_t = 3$)
2	PC$[t] = -0.1$: Decrease the transmit power once, ($\kappa_t = 1$)
3	PC$[t] = +0.1$: Increase the transmit power once, ($\kappa_t = 1$)
4	PC$[t] = +0.1$: Increase the transmit power thrice, i.e., total change is $+0.3$, ($\kappa_t = 3$)

- *States*: The set of states $\mathcal{S}$ is presented in Table II. We consider a total of five states based on the actions (outlined above) corresponding to changes in BLER, μ.

TABLE II: Set of states associated with change in BLER

State	Description
0	No change in BLER
1	Substantial increase in BLER
2	Moderate increase in BLER
3	Moderate decrease in BLER
4	Substantial decrease in BLER

- *Transition Probability*: The transition probabilities for each state-action pair, represented as $(p : \mathcal{S}\times\mathcal{A}) \to [0,1]$. Therefore, $p(s'|s,a)$ denotes the probability of transition to a new state s' from current state s with action a. It becomes very difficult to model all the transition probabilities using the traditional MDP for a real-world applications, such as cellular network with many UEs and gNBs. Instead, we adopt Q-learning algorithm which is model free well-defined reinforcement learning technique [13].
- *Rewards*: When the RL agent takes an action $a \in \mathcal{A}$ in state $s \in \mathcal{S}$ to transition to state $s' \in \mathcal{S}$, it gets a reward $r_{s,s',a}$. The obtained reward $r_{s,s',a}$ can be positive if the RL agent makes a move closer to the objective, else it is negative. We consider the following rewards:

$$r_{s,s',a} \triangleq \begin{cases} r_0, & \text{if } s' = s_0, \forall (s,a) \in \mathcal{S}\times\mathcal{A}, \\ r_1, & \text{if } s' = s_1, \forall (s,a) \in \mathcal{S}\times\mathcal{A}, \\ \vdots & \\ r_k, & \text{if } s' = s_k, \forall (s,a) \in \mathcal{S}\times\mathcal{A}. \end{cases} \quad (9)$$

As mentioned earlier, we adopt Q-learning algorithm which is model free and is based on probability distribution function, known as Q-value function $Q(s,a)$. It estimates the cumulative

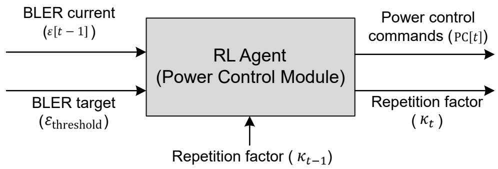

Fig. 1: Network parameters tuning by RL agent

expected reward for taking an action a in state s. The queue values $Q(s,a)$ are updated using:

$$Q_t(s,a) \leftarrow (1-\alpha)Q_{t-1}(s,a) + \alpha\left[r_{s,s',a} + \zeta \max_{a'} Q_{t-1}(s',a')\right]. \quad (10)$$

The symbols α and ζ in (10) denote the learning rate and discount factor, respectively. The values are as follows: α: $0 < \alpha < 1$, and ζ: $0 \leq \zeta < 1$.

B. Q-Learning-based Optimization Problem

The objective is to tune the power control commands (actions) to adjust the transmit power of each gNB for all the UEs connected to it for downlink transmission. Mathematically, we rewrite the optimization problem in (7) as follows:

$$\min_{a=[a_1,a_2,\ldots,a_\tau]} \sum_{t=1}^{\tau} \sum_{i=1}^{|N_{\text{UE}}|} P_{\text{tx}}^{(i)}[t], \quad (11)$$

subject to

$$\mu^{(i)}[t] \leq \mu^{(i)}_{\text{threshold}}, \forall i \in N_{\text{UE}}, \quad (12a)$$

$$P_{\text{tx}}^{(i)}[t] \leq P_{\text{tx}}^{\max}, \forall t \in \{1,\ldots,\tau\}, \quad (12b)$$

$$a_t \in \mathcal{A}, \forall t \in \{1,\ldots,\tau\}. \quad (12c)$$

Here in (12b), the transmit power is determined by:

$$P_{\text{tx}}^{(i)}[t] = \min\{P_{\text{tx}}^{\max}, P_{\text{tx}}^{(i)}[t-\Delta] + \kappa_t \text{PC}[t]\}, \quad (13)$$

where, $P_{\text{tx}}^{(i)}[t-\Delta]$ denotes the transmission power allocated for the last transmission. κ_t denotes the number of times the power control commands (actions) PC[t] executed, as mentioned in Table I. Figure 1 presents the schematic diagram of the network tuning parameters, where the RL agent tunes the power control commands to meet the target BLER. The repetition factor determines the number of times the power control command needs to be executed, as presented in Table I.

Algorithm 1 presents the proposed algorithm for closed loop power control. The time complexity associated with the power control algorithm is primarily associated with the state-action pair. Therefore, the time complexity is $O(|\mathcal{S}|\times|\mathcal{A}|)$ [14], where $|\mathcal{S}|$ and $\mathcal{A}$ denote the number of states and actions, respectively. The state space is exhaustive, which makes $|\mathcal{S}|$ fixed. Therefore, the effective time complexity of the power allocation algorithm is $O(|\mathcal{A}|)$.

Algorithm 1 Q-learning-based adaptive power control for downlink transmission

Inputs: Initial BLER: $\mu_0^{(i)}$; BLER threshold: $\mu^{(i)}_{\text{threshold}}$ for each user $i \in N_{\text{UE}}$; maximum transmission power: $P_{\text{tx}}^{\max}$; Set of states: $\mathcal{S}$; Set of power control commands (actions): $\mathcal{A}$
Output: The sequence of PC commands to achieve the BLER less than or equal to the threshold BLER in duration τ for a downlink transmission.
Note: We remove user index i for simplicity. The below steps are repeated for each user $i \in N_{\text{UE}}$.

```
The Q-table entries Q ∈ R^{|S|×|A|} are set to zero
Set time step t ← 0
Assign current BLER as: μ ← μ_0
Assign current state s ← 0
repeat
    Add time counter as: t ← t + 1
    Exploration rate update rule: ε ← max(ε × d, ε_min)
    Get reward from uniform distribution: r ∼ U(0, 1)
    if r ≤ ε then
        Pick an action a randomly from the action set A
                                        ▷ Explore new actions
    else
        Pick an action a = arg max_{a'} Q(s, a')
                                        ▷ Exploit optimal actions
    Execute action a and update P_tx^{(i)}[t] using (13)
    Get the reward r_{s,s',a} and observe next state s'
    Update Q-value table entry as:
        Q_t(s,a) ← (1 − α)Q_{t−1}(s,a)
                 +α [r_{s,s',a} + ζ max_{a'} Q_{t−1}(s',a')]
    Switch to new state: s ← s' and update μ
until μ ≤ μ_threshold or t ≥ τ
```

V. Performance Evaluation

A. Curve-Fit Parameters for Equation (3)

Table III shows the values of C_m and d_m for different modulation and coding schemes (MCS). To determine the values of C_m and d_m, we consider 32-byte payload for uRLLC applications and transport block size as 256 [2]. The simulation is conducted in MATLAB. Using the values of C_m and d_m, we obtain the SINR threshold values from the simulation.

B. Proposed Q-Learning-based Results

To evaluate the performance of the proposed scheme, we use the simulation parameters as presented in Table IV. The values of other communication parameters are considered based on the existing studies [4], [6], [8], [13]. To evaluate the performance of the proposed Q-learning-based approach, we consider the `fixed power allocation (FPA)` as the benchmark scheme. In `FPA`, the associated gNB allocates the same transmit power to each PRB, irrespective of the received SINR and the corresponding BLER for the downlink transmission. Furthermore, we present the results with MCS index 15 due to page limitations. We note that we observed the similar patterns for other MCSs. We consider retainability

TABLE III: Curve-fit parameters for different MCSs

MCS Index	**Modulation**	**Code Rate×1024**	C_m	d_m
0	QPSK	30	10.86×10^6	243.5
1	QPSK	40	7.78×10^6	179.3
2	QPSK	50	11.76×10^6	146.7
3	QPSK	64	6.41×10^6	112
4	QPSK	78	15.18×10^6	94.72
5	QPSK	99	12.58×10^6	74.15
6	QPSK	120	4.34×10^6	58.81
7	QPSK	157	1.24×10^7	46.57
8	QPSK	193	1.38×10^6	37.53
9	QPSK	251	6.66×10^6	26.55
10	QPSK	308	3.72×10^6	19.91
11	QPSK	379	2.67×10^6	15.33
12	QPSK	449	2.21×10^6	12.06
13	QPSK	526	1.28×10^6	9.275
14	QPSK	602	8.18×10^5	7.324
15	16-QAM	340	1.57×10^5	4.975
16	16-QAM	378	2.78×10^5	4.46
17	16-QAM	434	1.32×10^5	3.331
18	16-QAM	490	1.06×10^5	2.627
19	16-QAM	553	9.02×10^4	2.033
20	16-QAM	616	6.82×10^4	1.573
21	64-QAM	438	4.91×10^4	1.169
22	64-QAM	466	1.75×10^4	0.948
23	64-QAM	517	1.35×10^4	0.733
24	64-QAM	567	9.36×10^3	0.545
25	64-QAM	616	7.95×10^3	0.426
26	64-QAM	666	6.02×10^3	0.334
27	64-QAM	719	3.24×10^3	0.247
28	64-QAM	772	1.06×10^3	0.172

TABLE IV: Simulation Parameters

Parameter	**Value**
Bandwidth	20 MHz
Number of PRBs	100
Maximum transmit power ($P_{\text{tx}}^{\text{max}}$)	33 dBm
Discount factor (ζ)	0.995
Exploration rate (ϵ)	0.9
Learning rate (α)	0.2
Number of states ($\lvert\mathcal{S}\rvert$)	5
Number of actions ($\lvert\mathcal{A}\rvert$)	5

of BLER as the performance metric to show the efficacy of the proposed scheme. Where, retainability is mathematically defined as:

$$\text{Retainability} = 1 - \frac{1}{\tau|N_{\text{UE}}|}\sum_{t=1}^{\tau}\sum_{i=1}^{|N_{\text{UE}}|}\mathbf{1}_{\left\{\mu^{(i)}[t]>\mu^{(i)}_{\text{threshold}}\right\}} \tag{14}$$

where, $\mu^{(i)}[t]$ represents the BLER of i-th user at time t. The symbol $\mu^{(i)}_{\text{threshold}}$ is the threshold BLER to meet the reliability requirement of the underlying application. For the experiment, we consider the threshold BLER, $\mu^{(i)}_{\text{threshold}}$, between 10^{-1} and 10^{-4}. The expression $\mathbf{1}_{\{\cdot\}}$ represents the indicator function,

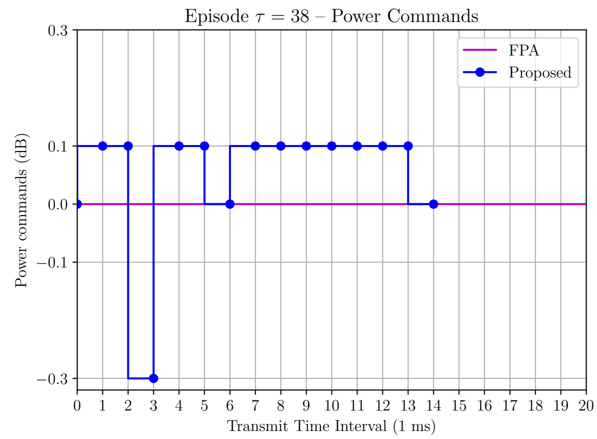

Fig. 2: Power control commands with time

which returns 1 if the condition is true, i.e., if the BLER is more than the threshold BLER to consider reliability, otherwise, it is zero. Table V shows the retainability in terms of BLER using `FPA` and the proposed scheme. We note that BLER is very sensitive with a small change in SINR (refer to (3)) due to power control commands. This leads to a lower retainability in terms of BLER as seen in the experiment. However, it is better than the fixed power allocation method.

TABLE V: Percentage of retainability

	FPA	**Proposed**
Score	55%	**60%**

In the following subsections, we discuss the experiment results on power control, average SINR, and average BLER with time to compare the proposed scheme with `FPA`.

C. Power Control

Figure 2 shows the change in power control commands using the proposed scheme and `FPA`. As depicted in the figure, `FPA` allocates fixed power irrespective of the received SINR and the associated BLER. Whereas in the proposed scheme, RL agent adjusts transmit power using power control commands, as shown in Figure 2. Intuitively, we say that the proposed scheme provides adaptive power allocation compared to the fixed power allocation scheme, `FPA`, as per the state of the downlink transmission link.

D. Average SINR

Figure 3 depicts the average SINR received by users at each time interval using the proposed scheme and `FPA`. As shown in the figure, the average SINR received by the users using the proposed scheme is higher than that of using `FPA` in most of the transmit time interval. Moreover, the average SINR is adjusted to meet the target BLER, as discussed in Section V-E.

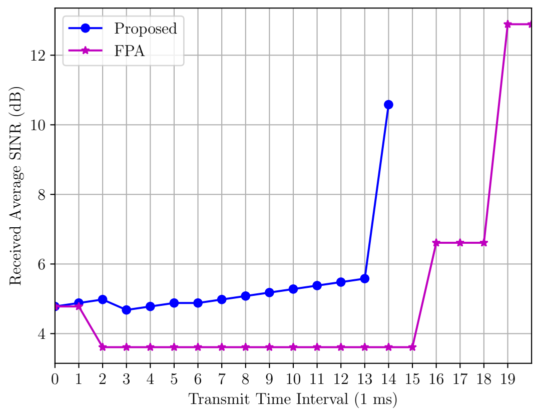

Fig. 3: Average SINR received by the users with time

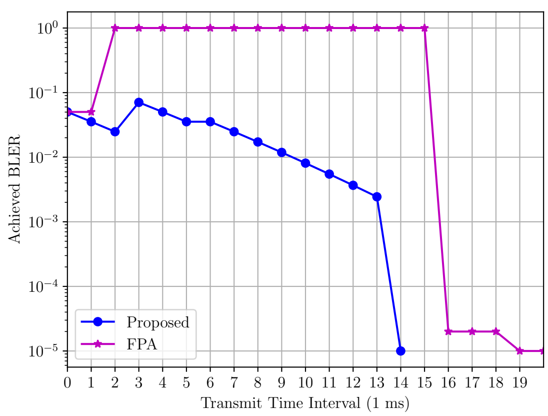

Fig. 4: Average BLER achieved by the users with time

E. Average BLER

Figure 4 depicts the average BLER achieved by the users. The proposed scheme achieves significantly lower BLER than that of using FPA. This is due to the adaptive power control by the RL agent, which adjusts the transmit power to achieve the required SINR to meet the target BLER. In contrast, FPA achieves a higher BLER, even as 10^0 in multiple transmit time intervals, due to fixed power allocation irrespective of the received SINR. We note that the transmitted information is useless at the user with a BLER 10^0. This is because of the fixed transmit power allocation, which is insufficient to meet the SINR threshold values. Therefore, FPA may not be suitable for applications with stringent reliability requirements.

VI. Conclusion

In this work, we proposed adaptive power allocation scheme for supporting stringent reliability requirements by uRLLC application in 5G and beyond 5G networks. To model the reliability of a communication link, we used block-error rate (BLER) for a downlink transmission and we derived the relation between BLER and SINR for the power allocation problem. To solve the non-convex optimization problem for power allocation, we used reinforcement learning, specifically, Q learning approach. The simulation was conducted considering well-established values for channel modeling and Q-learning approach. The results showed the efficacy of the proposed scheme over the fixed power allocation approach.

This work considered different communications impairments due to the circuit-level faults. In future, we plan to consider different environmental impairments in indoor and outdoor scenarios with the impairments considered in this work. We plan to use deep Q-learning approach to solve the power allocation problem with application-specific stringent reliability requirements and communication impairments. Furthermore, this work considered a fixed modulation and coding scheme (MCS) for transmission based on SINR and target BLER. However, we plan to consider different MCS for data transmission to effectively utilize the radio resources.

Acknowledgment

This work has been partially supported by the SRG research grant (grant number: SRG/2023/000569), funded by SERB, Govt. of India.

References

[1] J. Navarro-Ortiz, P. Romero-Diaz, S. Sendra, P. Ameigeiras, J. J. Ramos-Munoz, and J. M. Lopez-Soler, "A Survey on 5G Usage Scenarios and Traffic Models," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 905–929, 2020.

[2] 3GPP, "Physical layer procedures for data (3GPP TS 38.214 version 17.3.0 Release 17)," 3GPP, Tech. Rep. ETSI TS 138 214 V17.3.0, 2022.

[3] X. Zhang, M. Matthaiou, M. Coldrey, and E. Björnson, "Impact of Residual Transmit RF Impairments on Training-Based MIMO Systems," *IEEE Trans. Commun.*, vol. 63, no. 8, pp. 2899–2911, 2015.

[4] F. B. Mismar, J. Choi, and B. L. Evans, "A Framework for Automated Cellular Network Tuning With Reinforcement Learning," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 7152–7167, 2019.

[5] F. B. Mismar and B. L. Evans, "Q-Learning Algorithm for VoLTE Closed Loop Power Control in Indoor Small Cells," in *Asilomar Conference on Signals, Systems, and Computers*, 2018, pp. 1485–1489.

[6] M.-T. Nguyen, J. Song, S. Kwon, and S. Kim, "Power Allocation for Adaptive-Connectivity Wireless Networks Under Imperfect CSI," *IEEE Trans. Commun.*, vol. 71, no. 7, pp. 4328–4343, 2023.

[7] N. NaderiAlizadeh, M. Eisen, and A. Ribeiro, "Adaptive Wireless Power Allocation with Graph Neural Networks," in *IEEE ICASSP*, 2022, pp. 5213–5217.

[8] F. T. Al Rabee, A. Masadeh, S. Abdel-Razeq, and H. Bany Salameh, "Actor–Critic Reinforcement Learning for Throughput-Optimized Power Allocation in Energy Harvesting NOMA Relay-Assisted Networks," *IEEE Open J. Commun. Soc*, vol. 5, pp. 7941–7953, 2024.

[9] D.-D. Tran, S. K. Sharma, and S. Chatzinotas, "BLER-based Adaptive Q-learning for Efficient Random Access in NOMA-based mMTC Networks," in *IEEE VTC-Spring*, 2021, pp. 1–5.

[10] Z. Gao, B. Wen, L. Huang, C. Chen, and Z. Su, "Q-Learning-Based Power Control for LTE Enterprise Femtocell Networks," *IEEE Syst. J.*, vol. 11, no. 4, pp. 2699–2707, Dec. 2017.

[11] A. Rajanandini and T. Jaya, "Adaptive power allocation with energy efficiency in 5G multitier networks using a hybrid heuristic approach," *Sustain. Energy Technol. Assess.*, vol. 53, p. 102660, Oct. 2022.

[12] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[13] A. Barto and R. S. Sutton, *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press, 2018.

[14] S. Koenig and R. G. Simmons, "Complexity analysis of real-time reinforcement learning," in *Proc. of AAAI Conf. Artif. Intell.*, 1993, pp. 99–105.